

# Introduction to ROOT: part 1

Mirco Dorigo mirco.dorigo@ts.infn.it

LACD 2024-2025 March 19<sup>th</sup>, 2025



#### Contacts

- Researcher <u>@INFN Trieste</u>
- <u>mirco.dorigo@ts.infn.it</u> <u>mirco.dorigo@cern.ch</u>



- Worked in <u>CDF</u> (UniTS, 2009-2013) and <u>LHCb</u> (EPFL, CERN, 2013-2020)
- In <u>Belle II</u> since 2020
   <u>https://web.infn.it/Belle-II/index.php/our-research</u>

#### Class plan

#### Wed 19/03 Aula A, Ed A

Setup, basics commands and (very) little C++ tour

#### Fri 22/03 Aula B, Ed A

Reading and storing data (histograms, tuples)

#### Wed 27/03 Aula A, Ed A

Manipulating data (inspecting distributions, making selections, making graphs)

#### Wed 03/04 Aula B, Ed B

Fitting data

Main resource: <a href="https://moodle2.units.it/course/view.php?id=14880">https://moodle2.units.it/course/view.php?id=14880</a>

# https://root.cern.ch

- Open-source analysis framework with building blocks for:
  - ✓ Data processing
  - ✓ Data analysis
  - ✓ Data visualisation

CMS  $\sqrt{s} = 7 \text{ TeV}, L = 5.1 \text{ fb}^{-1} \sqrt{s} = 8 \text{ TeV}, L = 5.3 \text{ fb}^{-1}$ Ge∕ Unweighted S/(S+B) Weighted Events / 1.5 GeV <u>1500</u> Events / 1.000 1500 1000 120 130 m<sub>γγ</sub> (GeV) Data 500 S+B Fit B Fit Component ±1σ  $\pm 2 \sigma$ n 120 130 140 150 110 m<sub>γγ</sub> (GeV)

✓ Data storage

Physics Letters B 716 (2012) 30-61

- Widely use in high-energy physics (but not only):
   > 1EB of data in ROOT format at CERN, thousands of plots from ROOT in papers...
- Written mainly in C++ (bindings for Python available)



ROOT enables statistically sound scientific analyses and visualization of large amounts of data: today, more than 1 exabyte (1,000,000,000 gigabyte) are stored in ROOT files. The Higgs was found with ROOT!



As high-performance software, ROOT is written mainly in C++. You can use it on Linux, macOS, or Windows; it works out of the box. ROOT is open source: use it freely, modify it, contribute to it!

ROOT comes with an incredible C++ interpreter,

ideal for fast prototyping. Don't like C++? ROOT

integrates super-smoothly with Python thanks to

its unique dynamic and powerful Python  $\rightleftharpoons$  C++

binding. Or what about using ROOT in a Jupyter

notebook?

# Let's do a **real data analysis!**

- We will learn ROOT by doing an analysis using data from a real experiment, Belle II.
- Our goal is to see the signal peak of a rare *B* decay (branching fraction ~10<sup>-5</sup>):

$$B^0 \to K^+ \pi^-$$



- With ROOT we will optimise a selection to enhance our signal and measure its yield in our data.
- The study of this decay has been part of a real Belle II publications: <u>https://journals.aps.org/prd/pdf/10.1103/PhysRevD.109.012001</u> one of the three main authors was a Master student like you!

#### Root and C++

- <u>C++</u> is a coding language to program (writing instructions for your pc to execute).
- Here we won't learn C++: just very basic concepts to tell ROOT what to do.
- C++ is a compiled language: a compiler translates ASCII files with code into machine instructions. A compiler is gcc.
- ROOT comes with an interpreter (CLING), don't need to compile code to run it
  - it's not a C++ feature, its ROOT
  - CLING features just in time (JIT) compilation
  - CLNG provides an interactive C++ shell
- Very convenient: rapid prototype/check (drawback: learn sloppy C++...)

# Let's start ROOT

To start ROOT just type root in your shell



root [0]

- .q to quit ROOT
- . ? to obtain a list of command
- .!<command> (e.g. .!pwd) to access shell command
- Can start ROOT also with flags (eg. root -1).
  - -1 (do not show the root banner)
  - –b (batch mode, no graphics)
  - −q (run and quit)
- A few examples below, try man root for full list.

#### Using the prompt

As a simple calculator •

[mb-md-01:~ dorigo\$ root -1 [root [0] 2\*3 + 10 - 36 (int) -20 [root [1] 2\*3. (double) 6.000000 [root [2] pow(2,8) (double) 256.00000 [root [3] sqrt**(**144) (double) 12.000000

#### Accessing complex functions (via TMath library)



#### Using the prompt

#### To access ROOT classes



• Draw the function 1/(1-x)



10

#### Running a macro

• The prompt is powerful, but not convenient to (re)run several lines of code. Let's put them in a "macro", a bunch of lines of code in a ASCII file.

| voic myMacro()(                     |      |  |
|-------------------------------------|------|--|
| <pre>//several lines of codes</pre> | here |  |
| <pre>return; }</pre>                |      |  |

- Go back and put in a macro the example of the geometrical series.
- Note: the name of the macro must be the same of the function
- To run your macro, type root -1 myMacro.C, or

```
mb-md-01:~ dorigo$ root -1
root [0] .xemyMacro.C
Value after 20 iterations: 1.145475
root [1]
```

# Compiling a macro

- Not only JIT compilation, ACLIC can make libraries from your code
- Just load the macro adding a '+' at the end: .L myMacro.C+

```
[root [0] .L myMacro.C+
Info in <TMacOSXSystem::ACLiC>: creating shared library /Users/dorigo/./myMacro_C.so
In file included from input_line_12:6:
././myMacro.C:7:44: error: use of undeclared identifier 'pow'
for(int i=0; i<iterations; ++i) result += pow(variable,i);
././myMacro.C:16:2: error: use of undeclared identifier 'cout'
cout << "Value after " << N << " iterations: " << g_series(x,N) << endl;
././myMacro.C:16:69: error: use of undeclared identifier 'endl'
cout << "Value after " << N << " iterations: " << g_series(x,N) << endl;</pre>
```

• What's the problem?

#### Need to be C++ compliant

• Add some "headers"; make explicit the use of std (standard) library



Should be OK now

```
mb-md-01:~ dorigo$ nano myMacro.C
mb-md-01:~ dorigo$ root -1
root [0] .L myMacro.C++
Info in <TMacOSXSystem::ACLiC>: creating shared library /Users/dorigo/./myMacro_C.so
root [1] myMacro()
Value after 20 iterations: 1.145475
root [2]
```

# Going full C++

• ROOT libraries can be used to produce standalone compiled applications. Need to make our macro C++ standard code, by adding the main function



• Compile and run the binary example.

```
mb-md-01:~ dorigo$ g++ -o example myMacro.C
mb-md-01:~ dorigo$ ./example
Value after 20 iterations: 1.145475
mb-md-01:~ dorigo$
```

#### Language considerations

- Our code will be simple macros that can run on-the-fly, without compilation. We can afford being sloppy with the language...
- Anyway, a minimum knowledge of C++ basics is needed.
- Will have a look but you will mostly learn by copying examples.
   If you are completely unfamiliar, there are many good tutorials and guides on the web (e.g. <u>http://www.cplusplus.com</u>).
- Let's do a quick tour

#### Fundamental types



#### return;

| C++ Fundamental Types |               | Machine Independent Types |              |  |
|-----------------------|---------------|---------------------------|--------------|--|
| C++ type              | Size (bytes)  | ROOT types                | Size (bytes) |  |
| (unsigned)char        | 1             | (U)Char_t                 | 1            |  |
| (unsigned) short      | 2             | (U)Short_t                | 2            |  |
| (unsigned) int        | 2 or 4        | (U)Int_t                  | 4            |  |
| (unsigned)long        | 4 or 8        | (U)Long_t                 | 8            |  |
| float                 | 4             | Float_t                   | 4            |  |
| double                | 8 (>=4)       | Double_t                  | 8            |  |
| long double           | 16 (>=double) |                           |              |  |

#### Operators



Make actions on the variables, functions, output..

# (Some) operators

#### Arithmetic operators

| C++                                      | Purpose             |
|--|---------------------|
| x++                                      | Postincrement       |
| ++x                                      | Preincrement        |
| x  | Postdecrement       |
| x  | Predecrement        |
| +X                                       | Unary plus          |
| -X                                       | Unary minus         |
| x*y                                      | Multiply            |
| x/y                                      | Divide              |
| х%у                                      | Modulus             |
| x+y                                      | Add                 |
| х-у                                      | Subtract            |
| <pre>Pow(x,y) or TMath::Power(x,y)</pre> | Exp                 |
| х = у                                    | Assignment          |
| Х += У                                   | Updating assignment |
| X -=, *=, /=, %=,, Y                     |                     |

#### Logic/comparison operators

| C++             | ROOT extension |
|-----------------|----------------|
| false or O      | kFALSE         |
| true or nonzero | kTRUE          |
| !x              |                |
| х && у          |                |
| х    у          |                |
| х < у           |                |
| х <= у          |                |
| х > у           |                |
| х >= у          |                |
| х == у          |                |
| x != y          |                |

#### Loops et al. (statements)

- There are other types of loops (eg. while).
   They can be combined with other kind of statement, like
  if, if ... else ..., switch ... and so on
- We will see them with the examples throughout the lessons.

#### Functions

Very convenient to write functions in our macros

```
#include <math.h>
#include <iostream>
double g_series(double variable, int iterations){
 double result=0;
 for(int i=0; i<iterations; ++i) result += pow(variable,i);</pre>
 return result;
void myMacro(){
 double x = 0.127;
 int N = 20;
 std::cout << "Value after " << N << " iterations: " << g_series(x,N) << std::endl;</pre>
 return;
```

• Notice: myMacro() was used as a function in main in slide 9.

#### Functions – overloading

• Parameters are important. Can overload functions.

```
include <math.h>
#include <iostream>
double g_series(double variable){
 double result=0;
 for(int i=0; i<3; ++i) result += pow(variable,i);</pre>
 return result;
double g_series(double variable, int iterations){
 double result=0;
 for(int i=0; i<iterations; ++i) result += pow(variable,i);</pre>
 return result;
}
void myMacro(){
 double x = 0.127;
 int N = 20;
 std::cout << "Value after " << N << " iterations: " << g_series(x,N) << std::endl;</pre>
 std::cout << "Value after 3 fixed iterations: " << g_series(x) << std::endl;</pre>
 return;
```

#### Functions - overloading

• Parameters are important. Can overload functions.

```
include <math.h>
#include <iostream>
double g_series(double variable){
double result=0;
for(int i=0; i<3; ++i) result += pow(variable,i);</pre>
return result;
double g_series(double variable, int iter mb-md-01:~ dorigo$ root -1 myMacro.C
                                        root [0]
double result=0;
for(int i=0; i<iterations; ++i) result += Processing myMacro.C...</pre>
return result;
                                        Value after 20 iterations: 1.14548
                                         Value after 3 fixed iterations: 1.14313
void myMacro(){
                                        [root [1] .q
double x = 0.127;
int N = 20;
std::cout << "Value after " << N << " iterations: " << g_series(x,N) << std::endl;</pre>
std::cout << "Value after 3 fixed iterations: " << g_series(x) << std::endl;</pre>
```

#### Defining new types

• The first step to define new types is to create a structures to group elements (members)

#include <iostream> struct ComplexNumber{ double re; double im; }; void macro(){ ComplexNumber z; z.re = 1.;z.im = 3;std::cout << "real part: " << z.re << endl;</pre> std::cout << "imaginay part: " << z.im << endl;</pre>

A structure to define a new type, complex numbers

An object of the new type. Access the members re and im using a dot.

#### Defining new types

• The first step to define new types is to create a structures to group elements (members)

| <pre>#include <iostream></iostream></pre>  |    |
|--|----|
| <pre>struct ComplexNumber{</pre>   |    |
| double re;<br>double im;   |    |
| };   |    |
|  |    |
| <pre>void macro(){</pre>   |    |
| ComplexNumber z;<br>z re = 1 ·   |    |
| z.im = 3 ;   | [] |
| <pre>std::cout &lt;&lt; "real part: " &lt;&lt; z.re &lt;&lt; endl;<br/>std::cout &lt;&lt; "imaginay part: " &lt;&lt; z.im &lt;&lt; endl;</pre> | 1  |
|  |    |

| root  | [0]  | .x m | acro | o.C |  |
|-------|------|------|------|-----|--|
| real  | part | : of | z 1  |     |  |
| imagi | nay  | part | of   | z 3 |  |

#### Classes

Classes are structures on steroids: add functionalities (methods) #include <iostream> class ComplexNumber{ protected: double re; double im; public: class "constructor" ComplexNumber(double x, double y) { re = x; im= y; } double GetRe(){ return re; } double GetIm(){ return im; } Can define all operations void cPrint(){ std::cout << "Re: " << re << " " << "Im: " << im << std::endl;</pre> that you want with the } members of the class //can continue... //for instance, define sum, product, ... }; void macro(){ ComplexNumber z(3,4); Initialise an object std::cout << "real part of z " << z.GetRe() << std::endl;</pre> std::cout << "imaginay part of z " << z.GetIm() << std::endl;</pre> Access the methods z.cPrint(); with the dot.

#### Classes

• Classes are structures on steroids: add functionalities (methods)

```
#include <iostream>
class ComplexNumber{
double re;
ComplexNumber(double x, double y) { re = x; im= y; }
double GetRe(){ return re; }
double GetIm(){ return im; }
void cPrint(){
  std::cout << "Re: " << re << " " << "Im: " << im << std::endl;</pre>
/oid macro(){
ComplexNumber z(3,4);
std::cout << "real part of z " << z.GetRe() << std::endl;</pre>
std::cout << "imaginay part of z " << z.GetIm() << std::endl;</pre>
```

root [0] .x macro.C real part of z 3 imaginary part of z 4 Re: 3 Im: 4 root [1]

\_z.cPrint();

# Object oriented

- Classes have members (variables) and methods (functions)
- An instance of a class is an object, created by a special method, the constructor (can be overloaded).
- We can define very abstract classes, and then add derived classes that inherit from them to go more specific with what we need to do.



## Going back to ROOT

- ROOT is organised in classes: you will use objects and methods
- All classes begin with a "T" in ROOT (TGraph, TH1, TF1...)
- All methods begin with a capital letter (Draw(), GetX(), Derive()...)
- Classes inherited from more general (abstract) classes



28

#### Pointers

• Values are in memory, at a location (an address).



- & takes the address of value
- address now contains the memory-address of value
- \*address accesses the content



#### Pointers and objects

Can use pointers with objects: create with new



- Make explicit in code: ComplexNumber\* w = new ComplexNumber(3,2);
- Should need also a destructor to delete, but for simple classes like that the compiler takes care for us (important when you have pointers in the class, to free allocated memory).

#### Scope

- Every variables has a lifetime. It is defined only within a scope.
- It is determined by the  $\{ \dots \}$

```
#include <math.h>
#include <iostream>
void myMacro {{
    double x = 0.127;
    int N = 20;
    double g_series = 0;
    for(int i=0; i< N; ++i) g_series += pow(x,i);
    std::cout << "Value after " << N << " iterations: " << g_series << std::endl;
    return;
}</pre>
```

#### C++ overview wrap-up

- Done a very quick (and incomplete) tour of C++.
   This is NOT sufficient C++ for real-life.
- Sufficient to follow the course. We will do very simple coding (might not be really C++ kosher...).
- Important to understand basic concepts, such that you are not lost when navigating the ROOT class reference (eg. <u>https://root.cern.ch/doc/master/classTH1.html</u>)
- Writing macros will come with examples...

#### Our case analysis: setting the stage



#### The experiment

#### Collisions of (7 + 4) GeV electron-positron beams at $\sqrt{s} \simeq 10.6 \, {\rm GeV}$



Collisions of (7 + 4) GeV electron-positron beams at  $\sqrt{s} \simeq 10.6$  GeV



- Collisions of (7 + 4) GeV electron-positron beams at  $\sqrt{s} \simeq 10.5794$  GeV
- $e^+e^- \rightarrow h Beeble/BaBar~Befactoriese^+e(4S) \rightarrow Y (AS) \rightarrow BB$



• *B* mesons have a lifetime of  $\sim 1.5$  ps<sup>\*</sup>: we detect the decay products.



\*how much does it travel in the detector?



#### Our data



39

#### Our data



#### Some exercises

- Start ROOT. From the prompt look at the content of your folder, and look at the content of the folder above.
- Write a macro to compute the integral of  $x^2$  between -1 and 1. Don't use TF1, but compare your results with that of TF1.
- Compile the macro in ROOT (.L macro.C+) and run it.
- Explore the TF1 class. Look at the type 2, <u>expression using variable x</u> with parameters. Using this, write a normal Gaussian function in the range -5 and 5, set the mean to 0 and the std deviation to 1, and draw it. Get the value of the 2<sup>nd</sup> derivative at x = 0. Put all in a macro and run it.
- From the ROOT prompt: draw the Landau function.

#### Extra

- Some instructions (a few years old, but should still work) at this link: <u>https://www.unibo.it/sitoweb/gabriele.sirri2/contenuti-utili/df5f946d</u>
  - For Windows, follow the instructions under "run Ubuntu natively on Windows 10/11 without Virtual Machines."
  - For Mac: in addition to the instructions in the link, you can also use Homebrew (<u>https://brew.sh/index\_it</u>) or MacPort (https://www.macports.org/ install.php), see <u>https://root.cern/install/#macos-package-managers</u>
- ROOT page for installation, where you can find the link to pre-compiled binaries: <u>https://root.cern.ch/downloading-root</u>
- In case you need, a bash guide (get familiar with Sect. 1, 2 and 3): <u>https://swcarpentry.github.io/shell-novice/</u>





#### void func(){

root [1]

| <b>F1</b> f_sqrt( <b>"f","sqrt(x)",</b> 0,100); |  |
|---|--|
| out << f_sqrt.Eval(9) << endl;                  |  |
| _sqrt.Draw();                                   |  |
|   |  |
|   |  |
|   |  |
|   |  |
|   |  |

| Put it on a macro and run it.  | C1<br><u>File Edit View Options T</u> ools | <u>H</u> elp |
|--|--|--------------|
| <pre>void func(){</pre>  |  |              |
| <pre>TF1 f_sqrt("f","sqrt(x)",0,100);<br/>cout &lt;&lt; f_sqrt.Eval(9) &lt;&lt; endl;<br/>f_sqrt.Draw();</pre> |  |              |
| }  |  |              |
| root [0]   |  |              |
| Processing func.C<br>3   |  |              |
| Info in <tcanvas::makedefcanvas>: created defa</tcanvas::makedefcanvas>  | ult TCanvas with name c1                   |              |

# Stack and heap

- Text segment: code to be executed
- Initialised data segment: initialised global variable
- Uninitialised data segment: contains uninitialised global variables
- The stack: contains the frames, collections of all data associated with one subprogram call (one function)
- The heap: dynamic memory, requested with "new"



## Stack and heap



- Without the pointer, the function func() is in the stack, and its scope ends after closing the last "}". The program, made just by this function, ends and all variables inside the function are lost.
- "new" puts the object on the heap, escapes scope and the object survives.